

The use of Delphi and nominal group technique in nursing education: A review

Foth, Thomas ; Efstathiou, Nikolaos; Vanderspank-Wright, Brandi; Ufholz, Lee-Anne; Düttborn, Nadin; Zimansky, Manuel; Humphrey-Murto, Susan

DOI:

[10.1016/j.ijnurstu.2016.04.015](https://doi.org/10.1016/j.ijnurstu.2016.04.015)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Foth, T, Efstathiou, N, Vanderspank-Wright, B, Ufholz, L-A, Düttborn, N, Zimansky, M & Humphrey-Murto, S 2016, 'The use of Delphi and nominal group technique in nursing education: A review', *International Journal of Nursing Studies*, vol. 60, pp. 112-120. <https://doi.org/10.1016/j.ijnurstu.2016.04.015>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

Checked for eligibility: 03/05/2016

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

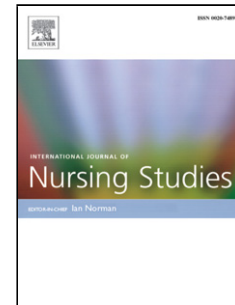
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Accepted Manuscript

Title: The Use of Delphi and Nominal Group Technique in Nursing Education: A Review

Author: Thomas Foth Nikolaos Efstathiou Brandi Vanderspank-Wright Lee-Anne Ufholz Nadin Dütthorn Manuel Zimansky Susan Humphrey-Murto



PII: S0020-7489(16)30035-9
DOI: <http://dx.doi.org/doi:10.1016/j.ijnurstu.2016.04.015>
Reference: NS 2743

To appear in:

Received date: 29-12-2015
Revised date: 29-2-2016
Accepted date: 25-4-2016

Please cite this article as: Foth, T., Efstathiou, N., Vanderspank-Wright, B., Ufholz, L.-A., Dütthorn, N., Zimansky, M., Humphrey-Murto, S., The Use of Delphi and Nominal Group Technique in Nursing Education: A Review, *International Journal of Nursing Studies* (2016), <http://dx.doi.org/10.1016/j.ijnurstu.2016.04.015>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Original Research Article:

The Use of Delphi and Nominal Group Technique in Nursing Education: A Review.

Word length (with references): 6537 words; (without references): 5600 words

Dr. Thomas Foth (corresponding author), RN, is assistant professor at the School of Nursing, Faculty of Health Sciences, University of Ottawa, Canada.
451 Smyth Road
Ottawa, Ontario
Canada
K1H 8M5
613.562.5800(8435) (P)
613.562.5443 (F)
tfoth@uottawa.ca

Dr. Nikolaos Efstathiou, RN, is Lecturer at the School of Nursing, Institute of Clinical Science, College of Medical and Dental Sciences, University of Birmingham, United Kingdom.

Dr. Brandi Vanderspank-Wright, RN, is assistant professor, School of Nursing, Faculty of Health Science, University of Ottawa, Ontario, Canada.

Lee-Anne Ufholz, MLIS, is a librarian at Wolters Kluwer Health based in Ottawa, Ontario, Canada.

Dr. Nadin Dütthorn, is assistant professor School of Health, department of Nursing and Health Care Education, University of Applied Science in Münster, Niedersachsen, Germany.

Manuel Zimansky, is doctoral student and lecturer, Department of Nursing Sciences, Faculty of Health Sciences, Osnabrück University, Niedersachsen, Germany.

Dr. Susan Humphrey-Murto is director of education research, Department of Medicine, the Ottawa Hospital, Ottawa, Ontario, Canada.

The Use of Delphi and Nominal Group Technique in Nursing Education: A Review

Abstract

Objectives:

Consensus methods are used by healthcare professionals and educators within nursing education because of their presumed capacity to extract the profession's "collective knowledge" which is often considered tacit knowledge that is difficult to verbalize and to formalize. Since their emergence, consensus methods have been criticized and their rigour has been questioned. Our study focuses on the use of consensus methods in nursing education and seeks to explore how extensively consensus methods are used, the types of consensus methods employed, the purpose of the research and how standardised the application of the methods is.

Design and Data sources:

A systematic approach was employed to identify articles reporting the use of consensus methods in nursing education. The search strategy included keyword search in five electronic databases [Medline (Ovid), Embase (Ovid), AMED (Ovid), ERIC (Ovid) and CINAHL (EBSCO)] for the period 2004-2014. We included articles published in English, French, German and Greek discussing the use of consensus methods in nursing education or in the context of identifying competencies.

Review Method:

A standardized extraction form was developed using an iterative process with results from the search. General descriptors such as type of journal, nursing speciality, type of educational issue addressed, method used, geographic scope were recorded. Features reflecting methodology such as number, selection and composition of panel participants, number of rounds, response rates, definition of consensus, and feedback were recorded.

Results:

1230 articles were screened resulting in 101 included studies. The Delphi was used in 88.2% of studies. Most were reported in nursing journals (63.4%). The most common purpose to use these methods was defining competencies, curriculum development and renewal, and assessment. Remarkably, both standardization and reporting of consensus methods was noted to be generally poor. Areas where the methodology appeared weak included: preparation of the initial questionnaire; the selection and description of participants; number of rounds and number of participants remaining after each round; formal feedback of group ratings; definitions of consensus and a priori definition of numbers of rounds; and modifications to the methodology.

Conclusions:

The findings of this study are concerning if interpreted within the context of the structural critiques because our findings lend support to these critiques. If consensus methods should continue being used to inform best practices in nursing education, they must be rigorous in design.

Keywords: Delphi technique, Nominal Group Technique, Consensus methods, nursing education

Accepted Manuscript

Contribution of the Paper

What is already known about the topic?

- Consensus group research methods are widely used in nursing research and in many other fields of research.
- Consensus group research methods derive quantitative estimates through qualitative approaches that should follow strict methodological guidelines.
- From their implementation into different fields of research since the 1950s, consensus research methods have been criticized for systematic shortfalls.

What this paper adds

- We focus on the use of consensus methods in the nursing education literature and analyze the main areas in which these methods are used.
- We discuss and support some of the critiques that question the validity of the method.
- If consensus methods are to be used to inform best education practice, they must be planned and executed rigorously.

1 Background

Consensus group methods such as the Delphi and Nominal Group Technique have been used since the 1950s as ways to collect opinions of a wide range of experts and to develop consensus between them. The Delphi technique was originally developed by the RAND Corporation in California and used as a method to identify potential key nuclear targets in the United States from an Union of Soviet Socialist Republics perspective (Campbell and Cantrill, 2001). However, the ambitions of the proponents of the Delphi technique went far beyond the political field. For authors like Helmer and Rescher from the RAND Corporation, this technique enabled scientific predictions and explanations in areas in which no empiric evidence, existed. If predictions in these cases could be achieved “correctly and in a systematic and reasoned way” they had to be classified as scientific (Helmer and Rescher, 1959, 25).

Since its introduction, the Delphi technique has been used for multiple purposes. Authors differentiate between the ‘classical Delphi’ used to determine facts, the ‘policy Delphi’ used to create ideas, and the ‘decision Delphi’ used to achieve decisions (Crisp et al., 1997). Over time the Delphi technique itself has been modified and other forms of consensus seeking methods have been developed (Murphy et al., 1998). Another commonly used method is Nominal Group Technique. The so-named RAND is a hybrid of the two. What these various approaches have in common is the use of a structured method for evaluating the degree to which experts agree about a particular issue, the assumption being that accurate and reliable assessment can best be achieved by

consulting a panel of experts, and accepting group consensus (Campbell and Cantrill, 2001, Tammela, 2013). Consensus methods or techniques are supposed to derive quantitative estimates through qualitative approaches (Jones and Hunter, 1995) by using processes which are characterized by several common features including anonymity, iteration, controlled feedback, statistical group response and structured interaction (Jones and Hunter, 1995, Murphy et al., 1998).

Consensus group methods are extensively used in many fields including business as well as healthcare research including medicine, nursing, health services research, training and education (Campbell and Cantrill, 2001, Jones and Hunter, 1995, Murphy et al., 1998, Tammela, 2013). Consensus group methods are used within the context of healthcare education because of their presumed capacity to extract the profession's' "collective knowledge" which is often described as tacit knowledge that is both difficult to verbalize and to formalize (Stewart, 2001). Consensus methods help to synthesize knowledge by including information that cannot be obtained through statistical methods (Jones and Hunter, 1995) and they are thought to enable decision making especially in "grey areas" of medicine (Naylor, 1995) and nursing that are not supported by evidence gained through clinical trials or other research.

1.1 The methods - benefits and limitations: The Delphi technique

Delphi technique generally involves the following stages: identifying a research problem, completing a literature review, development of a questionnaire of statements, conducting anonymous iterative postal or email questionnaire rounds in which the experts are asked to rate or rank the statements and determining whether they agree or disagree with the statements, individual and group feedback between rounds, consensus building and

summary of findings. The iterative process is continued until the greatest level of consensus is reached or a pre-determined number of rounds are completed. Participants do not meet face to face or interact directly (Boulkedid et al., 2011, Murphy et al., 1998, Sinha et al., 2011).

Benefits of the Delphi technique include the potential inclusion of large numbers of participants who are geographically dispersed and are from diverse areas of expertise (Jairath and Weinstein, 1994). Delphi technique enables academic expertise to be combined with practitioners' perspectives and experiences (Trevelyan, 2015) and may include patients, stakeholders and other "lay expertise" (Padgett et al., 2014). Through its anonymous and structured process, Delphi technique attempts to avoid some of the disadvantages of decision making processes in traditional group meetings such as the risk of meetings being dominated by one individual or influenced by coalitions between group members (Jairath and Weinstein, 1994).

However, since its emergence, Delphi technique has been subject to critique. Some of these criticisms include questioning the validity and reliability of the methodology; others criticize particular aspects of the method. One of the more fundamental critiques concerns the question of what some scholars called "imposed consent" (Haug, 2015). From this regard, the Delphi is criticized because it does not provide a space for important discussions and debates and provides no opportunity to discuss and resolve differences of opinion but rather bypasses disagreement by assuming it does not exist (Woudenberg, 1991). Other critics stated that the reasons why experts may rate an item low are not clear - is it because the expert assumes that no evidence

exists to support a specific statement or because he or she believes that the statement is not important (Goodman, 1987). Others concluded that the Delphi technique's consensus provides "bland generalities that represent the lowest common denominator of a debate – the only points on which the experts can wholeheartedly agree" (Fink et al., 1984, 3, Kea and Sun, 2015). Consensus in these situations may be trivial and only "stating well-established facts" rather than lead to innovations (Rennie, 1981, 666). Some critics go even as far as stating that consensus methods only "pool ignorance" (Greco et al., 2015, Scott, 1991, 1068). Due to the anonymity of the process, it has been argued that experts are not accountable for the views they express and the judgements they make (Fink et al., 1984). Related to the latter, is the question of how consensus should be measured and conceptualized. Rowe et al. (1991) emphasized that after the second or third Delphi round, little change in experts' opinions occurred which provided a justification to limit the Delphi to a predetermined number of rounds and to consider the results as a successful consensus (Granovskiy et al., 2015, Rowe et al., 1991). The same is true for the determination of a numerical value (e.g. standard deviation) as an indicator of consensus (Dajani et al., 1979, von der Gracht, 2012).

Another major critique concerns the definition and use of so called 'experts' in the Delphi technique. This critique extends to decisions regarding the size of a panel and the specific criteria/characteristics that determine if an expert will be included or excluded (Rowe et al., 1991). Authors like Sackman (1974) have questioned the assumption that the quality of expert opinion is superior to the opinions obtained from informed individuals (Sackman, 1974). Crisp et al. (1997) emphasized that authors

should provide their rationale for the type of panel they choose and concluded that “decisions concerning panel members are by no means as straightforward as they appear to be when represented in the literature” (Crisp et al., 1997, 117). Others emphasize that the composition of the panel can affect the results (Jones and Hunter, 1995, Scott, 1991).

Additionally, the feedback mechanisms of the Delphi technique have been questioned. According to some authors the feedback mechanism has a rather manipulating effect leading participants towards conformity instead of towards true consensus. According to these critics, once the panellists know how the others have rated each statement, they no longer independently rate each item. The Delphi technique “centralizes opinions” and if the feedback measures consensus by using mean or median variation of statements “it is not only insufficient but also potentially misleading” (Goodman, 1987, 733).

Aside from these more fundamental critiques, the number of modifications to the Delphi technique has led to considerable confusion (Crisp et al., 1997). For example, in the “Modified Delphi” a face-to-face component may be included (Murphy et al., 1998). Furthermore, a significant challenge associated with the method is that little guidance exists to help researchers undertake consensus methods of data collection and studies using these techniques often lack methodological rigour or vary in how they are implemented (Boulkedid et al., 2011, Diamond et al., 2014, McKenna, 1994, Sinha et al., 2011).

1.2 The Nominal Group Technique (NGT) and RAND

The NGT is a structured face-to-face interaction. The research team formulate a nominal question (questions with non-ordered response categories) and then present it to the participants when they meet. Initially, each participant records his or her ideas independently and privately. Participants then share their individual ideas with the group in a 'round-robin' format, with each participant sharing one item from their list. These are recorded by a facilitator who will then lead a group discussion of each idea. Individuals will then vote privately on the items and the results are provided to the group in aggregate (anonymously). Further discussion and voting may take place. Potential benefits include significant idea generation and the fact that discussion and debate may take place - even though it may be in a limited and pre-structured format (Murphy et al., 1998). Limitations of this consensus method include a smaller number of participants, and the potential for dominant participants to unduly influence the group.

The RAND is a hybrid and begins like the Delphi by identifying a research problem, completing a literature search, and developing of a questionnaire of statements that is mailed out. The next step, however, involves a face-to-face meeting (Crisp et al., 1997).

Many of the afore-mentioned critiques of Delphi technique also apply to NGT and RAND, for example, the aspect of imposed consent, defining and measuring consensus, definition of what characterizes an expert and how to compose the panel. Even though consensus methods are widely used in research, the methods are neither standardized nor are they consistently described and used (Crisp et al., 1997). Findings demonstrate lack of

consistency in nomenclature, poor documentation of methodology and lack of rigour (Crisp et al., 1997, Hasson et al., 2000). Several authors have noted poor agreement in defining Delphi with an ever-expanding list of modifications without validation thereby compromising the validity of the original approach (McKenna, 1994). In light of the fundamental concerns that question the validity and reliability of these techniques, the use of a sound methodology and transparency for every step in the consensus seeking process is paramount. However, an examination of the literature reveals a remarkable lack of detail provided by researchers in relation to processes used for generating items, reducing items or deciding on feedback (Crisp et al., 1997). Others have expressed concern regarding the lack of methodological scrutiny of consensus methods (Boulkedid et al., 2011, Campbell and Cantrill, 2001, Diamond et al., 2014, Keeney, 2001). A review in medical education has demonstrated similar concerns. For example, the quality of reporting was variable with just over half of the articles reporting on the number of respondents after each round, and well under half indicating what background information was provided to participants, whether private decisions were collected, if formal feedback of group ratings was provided, and only a third defined consensus a priori (Humphrey-Murto S, 2014).

Our study focuses on the nursing education literature and seeks to explore the extent to which consensus methods are used in nursing education research, the types of consensus methods used, the purpose or intent of the research and how standardised the application of the methods is. Our definition of nursing education was not restricted to academic nursing education but included, nursing education in clinical or other practice

settings as well as education in the context of professional development. Following the findings of our review, we discuss the strengths and weaknesses of the methods as used in nursing education and we identify the impact that these findings may have, including the broader political influence for nursing education.

2 Method

A pilot search was conducted in Ovid Medline to determine whether or not the search strategy was robust. A pilot keyword search began in October 2014 and was completed in November 2014 in the databases Ovid Medline and CINAHL focusing on nursing education and the following terms: Delphi, RAND, nominal group, consensus group methods - all with no date limit initially. We used an iterative process to identify the search strategy that would most appropriately identify the articles using consensus methods. We searched the databases Medline (Ovid), Embase (Ovid), AMED (Ovid), ERIC (Ovid) and CINAHL (EBSCO) which retrieved 1596 results. There were 1203 results remaining after duplicates were removed (number of duplicates 366).

The titles and abstracts were screened by two reviewers to identify relevant studies (LAU, TF). We included articles published in English, French, German and Greek that were using consensus methods in education, or in the context of identifying competencies. The concept education – both academic education and education in practice settings - was defined broadly and included the evaluation, development or design of specific teaching/learning units, curriculum development, renewal and evaluation, as well as the development and evaluation of specific teaching/learning interventions. We excluded studies reported only in abstract form, editorials, doctoral dissertations, methodological studies, comments and duplicate publications. For logistical

reasons, we restricted our search to 10 years (2004-2014). Based on previous research (Humphrey-Murto et al., 2014) this was deemed to provide a sufficient sample size for analysis.

A standardized extraction form was developed using an iterative process with results from the search. The form was a revised version of an extraction form from a study about the use of consensus methods in medical education (Humphrey-Murto S, 2014). The research team evaluated the articles and extracted data independently (TF, NE, MZ, LAU, ND, BVW). Each team member reviewed 10-25 papers. In order to check for interrater agreement, twenty-two articles were randomly allocated to different research team members for data extraction in order to check for inter-rater agreement. We only included those categories in our analysis that had achieved at least 80% or more inter-rater agreement in the randomly allocated twenty-two articles. The data recorded included a range of variables, such as the name of the journal, date of publication, nursing speciality, type of educational issue addressed and the content for consensus group being used for readership. Furthermore, consensus methods' specific features were recorded which are indicative of the rigour associated with the use of consensus methods (Boulkedid et al., 2011; Goodman, 1987; Hasson et al., 2000) (Table 1).

Table 1: Consensus methods features data extraction

Type of Consensus Method	Delphi, Nominal Group Technique, RAND, other
Participants	Number of participants
	Composition of panels
	How participants were chosen
Process of study	Type of 1 st round (generation of items, pre-determined items)
	Number of rounds (pre-determined or not)

	Response rates for each round
	Method used to send questionnaires
	Rating of items
	Definition of consensus
	Group and individual feedback
	Geographic scope

A member of the team (NE) collected all the data and analysis was performed using SPSS (v21). The analysis included computation of the number and percentage of articles for categorical variables and medians for continuous variables. For each characteristic, percentages were computed using as denominator the total number of articles reporting that characteristic.

3 Results

Following application of exclusion and inclusion criteria 101 articles were included in the analysis (Diagram 1). Most included articles were in English (97/101) and of the 101 consensus methods articles related to nursing education, 89/101 (88.2%) studies had described using a Delphi technique, 11/101 (10.9%) a Nominal Group Technique and one study was identified as using both. 64 of the 101 articles (63.4%) were published in nursing journals; usually a journal related to the speciality of nursing where the educational research was undertaken. Only 16/101 (15.8%) of the studies were published in healthcare education related journals. 21/101 (20.8%) of the articles were published in other health related journals such as health informatics, nursing management and administration (Table 2).

Table 2: Characteristics of the 101 selected articles.

Journal type	N (%)
Nursing Speciality	34 (33.6%)
General Nursing	30 (29,7%)

Healthcare Education	16 (15.8%)
Other	21 (20.8%)
Countries of origin	
United States	49 (48.5%)
United Kingdom	39 (38.6%)
Australia	9 (8.9%)
Others	4 (4%)
Range of years when studies were published	
2004 - 2009	43 (42.9%)
2010 – 2014	58 (57.1%)
Purpose for consensus group being used	
Continuing professional development	43 (42.6%)
Undergraduate nursing	17 (16.8%)
Postgraduate nursing	14 (13.9%)
Other	15(14.8%)
Not clear	12(11.9%)

Consensus methods were used in a variety of specialities with critical care nursing attracting most publications (10/101, 9.9%), followed by emergency nursing (5/101, 5%) and psychiatric nursing (4/101, 4%). The remaining articles (82/101, 81.1%) included specialities such as public health, rheumatology, nephrology, dermatology or general undergraduate nursing. Defining, developing or assessment of nursing competencies were the most common reasons (42/101, 41.6%) for using consensus methods, with 39/42 (93%) of the competencies related articles seeking consensus on defining competencies. Another 19/101 (18.8%) articles reported studies on development or renewal of curricula and 14/101 (13.9%) articles reported research on learning assessment methods or tools. For 43/101 (42.6%) of the articles, the topic was related to nurses in practice (post-registration), 17/101 (16.8%) studies were related to undergraduate nursing and 14/101 (13.9%) were related to post graduate nursing (Table 2). Interestingly, 12/101 (11.9%) of

the articles were not clear about what the purpose of the study was (i.e. why was a consensus group method used). For continuing professional development and postgraduate nursing, the most common topic seeking consensus was for defining competencies (18/43, 41.9% and 6/14, 42.8% respectively), while in an undergraduate nursing context, it was for renewal of curricula and assessment of competencies (5/17, 29.4% and 6/17, 35.2% respectively).

Participants in the studies varied, with 21/101 (20.8%) articles identifying nurses as participants, 19/101 (18.8%) studies having an intra-professional panel consisting of various nursing specialities, 16/101 (15.8%) having an inter-professional panel and 15/101 (14.9%) having only educators in their panel. In 8/101 (8%) of the articles, students were identified as participants (alone or members of a panel). In only 2/101 (2%) of the articles, patients were included as participants. In 6/101 (6%) articles, participants were reported as 'experts'. Out of the 101 studies, 87 were seeking consensus on a national (71.3%) or local (14.9%) educational nursing issue. A further 9 (8.9%) studies were seeking to address an international educational issue. For 5/101 (5%) of the articles, the study did not describe the scope of the study, i.e. local, national or international. Nearly half of the national and international level studies were seeking consensus for nursing competencies.

3.1 Delphi studies

Of the 89 articles defining the Delphi Technique as their method, 62 (69.7%) were reported as classical Delphi studies, 22 (24.7%) as modified Delphi studies and 5 (5.6%) used Delphi Technique combined with other methods (e.g. interviews, etc.). The majority of the Delphi studies (67/89, 75.3%) were conducted either via mail or electronically. Six

studies (6/89, 6.7%) reported a combination of mail and face-to-face conduct (these studies could be classified as RAND, however the authors did not identify them so) and for 15/89 (16.8%) studies the researchers did not report the method explicitly. One Delphi study (1/89, 1.1%) was conducted face to face.

Participants in 23/89 (25.8%) of the Delphi studies were chosen because they were considered by the researchers as experts in their speciality. In 13/89 (14.6%) of the Delphi studies participants were chosen because they were members of an organisation, in 9/89 (10.1%) the criterion was years of experience and in another 9/89 (10.1%) participants were recommended to the researchers. For 5/89 (5.6%) of the Delphi studies, participants were selected randomly. For 20/89 (22.6%) studies, a combination of the above was reported. For the remaining 10/89 (11.2%) studies, there was no clear explanation regarding the selection of the panel participants.

For half 47/89 (52.8%) of the Delphi studies, the initial questionnaire was prepared by the researchers, usually following a literature review or focus groups which in half of the cases were not described in detail. Twenty three (23/89, 25.8%) articles reported that the purpose of the first round was for item generation and 16/89 (18%) articles reported using a combination of a predefined questionnaire with item generation through consensus (Table 3). For a small number of articles 3/89 (3.4%) the researchers did not describe how the items were developed for the initial questionnaire. Two or three rounds were equally the most common form 78/89 (87.6%) of Delphi studies and only 3/89 (3.4%) reported a fourth round. For the remaining 8/89 (9%) studies, the number of rounds used was not clear (Table 3).

Table 3: Characteristics of Delphi procedure

Number of rounds	n (%)
<i>Delphi</i>	62 (69.6%)
<i>Median</i>	3
<i>Modified Delphi</i>	22 (24.7%)
<i>Median</i>	2
<i>Delphi & other</i>	5 (5.6%)
<i>Median</i>	2
Round 1	
<i>Delphi</i>	
<i>Pre-determined</i>	30 (48.4%)
<i>Item generation</i>	20 (32.2%)
<i>Both</i>	9 (14.6%)
<i>Not described</i>	3 (4.8%)
<i>Modified Delphi</i>	
<i>Pre-determined</i>	13 (59.1%)
<i>Item generation</i>	2 (9.1%)
<i>Both</i>	7 (31.8%)
<i>Not described</i>	0
<i>Delphi & other</i>	
<i>Pre-determined</i>	2 (40%)
<i>Item generation</i>	1 (20%)
<i>Both</i>	2 (40%)
Group feedback between rounds	
<i>Delphi</i>	
<i>Yes</i>	24 (38.7%)
<i>No</i>	2 (3.2%)
<i>Not described</i>	36 (58.1%)
<i>Modified Delphi</i>	
<i>Yes</i>	7 (31.8%)
<i>No</i>	5 (22.7%)
<i>Not described</i>	10 (45.4%)
<i>Delphi & other</i>	
<i>Yes</i>	1 (20%)
<i>No</i>	1 (20%)
<i>Not described</i>	3 (60%)
Pre-defined consensus	

<i>Delphi</i>	
<i>Yes</i>	32 (51.6%)
<i>No</i>	7 (11.3%)
<i>Not described</i>	23 (37.1 %)
<i>Modified Delphi</i>	
<i>Yes</i>	7 (31.8%)
<i>No</i>	6 (27.3%)
<i>Not described</i>	9 (40.9%)
<i>Delphi & other</i>	
<i>Yes</i>	1 (20%)
<i>No</i>	1 (20%)
<i>Not described</i>	3 (60%)

The number of participants for each round was reported by most researchers, but not always clearly. The number of participants ranged from round to round in some studies, rendering the calculation of response rates impossible. Attrition was higher in studies with a large number of participants, for example, a study of 1508 participants, had 16.1% response rate in round three.

With regards to the provision of feedback to participants between rounds, only 32/89 (35.9%) studies reported providing group feedback and 8/89 (9%) did not provide any feedback. Forty nine (49/89, 55.1%) articles did not describe the provision of feedback in between rounds.

Consensus was pre-defined by researchers in 40/89 studies (44.9%). It was predefined in half of the claimed classical Delphi studies (32/62, 51.6%) but only in one third of the modified Delphi studies (7/22, 31.8%). In 14/89 (15.7) of the studies, researchers reported that they did not define consensus prior to data analysis. For 35/89 (39.3%) studies there was no mention of predefined consensus. For 35/89 (39.3%) of the studies, consensus was described as percentage of agreement for an item, usually 60%

agreement or higher (median=75%). Other forms of agreement included acceptance of items that were rated at the upper extremes of the Likert scales used (for example items scored only as 4 and 5 on a 5 point Likert type scale). The most common form type of rating forms, were Likert type scales ranging from 3 to 10 points (importance, agreement, essential, relevance).

3.2 Nominal Group Technique

Ten of the 11 studies that used an NGT approach were conducted face-to-face and one a combination of mail and face-to-face. Participants were selected because they were considered experts, members of organisations, randomly assigned or no explanation was given in equal numbers of the articles.

In more than half (6/11, 54.5%) of the NGT studies the first round was used for generation of items, in two studies it was predetermined and in one study both predetermined and new item generation. For one article the researchers had not described how the first round was developed. The number of participants ranged from 4 to 121. For more than half of NGT (6/11, 54.5%), group feedback was not required due to the way the technique was used. For the remaining NGT studies, in 2 studies feedback was given, in 2 this was not described and in 1 study it was explicit that feedback was not given.

3.3 Delphi and NGT

Finally, the study using both a Delphi approach and NGT claimed in the abstract that the researchers employed qualitative methodology including a modified Delphi and nominal group technique, however, aspects related to these methods (number of rounds, response rate between rounds, level of consensus) were not described in detail.

4 Discussion

Our review demonstrates that consensus methods are widely used in nursing education particularly with respect to defining, developing, and assessing nursing competencies; curriculum development or renewal; and developing assessment tools. This result is not surprising considering that consensus methods are a means to extract non-formalized knowledge of the profession (Naylor, 1995, Stewart, 2001), and are similar to the medical education literature (Humphrey-Murto S, 2014).

The most important result of our study related to the methodologies used, is that both standardization and reporting of consensus methods is generally poor. There were several areas where the methodology appeared weak. These included: preparation of the initial questionnaire; the selection and description of participants; number of rounds and number of participants remaining after each round; formal feedback about group and individual ratings; definitions of consensus and a priori definition of numbers of rounds; and modifications to the methodology. Very few studies provided the number of invited participants for the initial round of the consensus methods. Similarly, few studies clearly reported response rates for subsequent rounds. Furthermore, few studies adequately described the participants in order to allow readers to make judgments about the quality of the “panel of experts”. This echoes work done by Boulkedid et al. (2011).

Additionally, only a small number of studies reported about providing feedback (individual or group) to participants, a feature which is regarded as major strength of consensus methods (Keeney, 2001, Murphy et al., 1998). Consensus within Delphi technique and NGT has been discussed extensively by scholars (Hasson et al., 2000, Keeney, 2001, Murphy et al., 1998, Sinha et al., 2011), however, researchers still fail to

address this issue successfully in their studies. Consensus was pre-defined only in less than half of the Delphi studies in our review and it occurred more commonly in classical Delphi studies.

Our findings correspond with findings about the use of consensus methods in other research areas. Two recent reviews have further highlighted the issues found in our review. Boulkedid's (2011) systematic review, which included 80 studies adopting the Delphi as the method to select quality indicators, found that reporting was poor. Of those reviewed, 39% of studies reported response rates for all rounds, 60% described the feedback provided and 77% properly explained how consensus was achieved. In a similar systematic review of 15 studies using Delphi designed to select outcome measures for clinical research, Sinha (2011) found that the following were poorly reported: information provided to the participants at the start of the Delphi, the information fed back to participants after each round, the level of anonymity, the attrition rates and the list of outcomes after each round (see, also, Keeney and McKenna, 2001).

Further to the common weaknesses described in our review and other consensus methods reviews, we identified important limitations related to their use in nursing education. For example, in relation to the participants in the studies we reviewed, we identified that most "expert panels" were composed either exclusively of nurses "working at the bedside" in combinations with nursing administrators, stakeholders from nursing associations, other healthcare professionals, or in few occasions, combinations of these specialists with clinical educators. In the majority of instances there was either minimal or no input from nursing educators/faculty in academic settings. Therefore, from a

critical perspective, it could be argued that the panel configurations we identified in the literature for defining and agreeing upon competencies and curricula are not necessarily ideal. Although contributions from clinical practice *are essential* in informing curricula and defining competencies (Black et al., 2008, College of Nurses of, 2014) there is a potential risk that consensus methods' panels, as identified in our review, may create nursing education that is entirely "practice" driven and reductionist, in that technical tasks are emphasized and where other intellectual skills (for example critical thinking and reflection) and other essential aspects of university education (including the application of research and theory and existential experiences related to nursing practice) are minimized (Carper, 1978, Evans, 1995). Therefore, the design of nursing education research utilising consensus methods must ensure representation of all stakeholders' (both frontline and academic) views and allow democratic decisions to be taken based on the outcomes of the research. Furthermore, it is imperative to use the strengths of consensus methods such as wider geographical representation (Murphy et al., 1998), since in our review we identified that researchers rarely exploited the possibility to integrate different opinions from national, let alone international contexts.

Although this study was not designed to systematically seek out political agendas, some became evident in the course of the study. We found that consensus methods appeared to be replacing free expression of different political positions by forcing decisions to be based on availability of resources, efficiency and efficacy (depending on the question the participants are expected to answer and subsequently agree on). One example for this displacement is the article by Fischer (2015) in which the author

proposes to use the Delphi technique to define technical efficiency and efficacy indicators in order to answer the political and highly controversial question regarding the sustainability of healthcare systems (Fischer, 2015). It was also noted that opposing or controversial views could be lost in consensus methods and the final consensus could be very conservative or reduced to basic technical terms. One example for how controversies are transformed into consensus is provided in an article from medicine (Nguyen-Lu et al., 2015). Controversies between different healthcare professionals on a ward regarding the interpretation of evidence and patient care were transformed into “consensus building” using Delphi technique thereby transforming conflicts (how care was defined by different professionals) into technical terms of patient care management.

Although we did not explore whether the findings from the consensus methods articles we reviewed were actually implemented in nursing education, it is worth noting that caution should be given in the direct application of findings from any consensus method due to the weaknesses identified. Furthermore, the use of “experts”, common in the studies we reviewed, raises questions about the level of evidence produced. Expert opinion is considered as the lowest level in the hierarchy of available evidence (DiCenso et al., 2005, Sackett, 1997, Sackett et al., 2000), and despite assurances by proponents of consensus methods who predicated that “two heads are better than one, or...n heads are better than one (Dalkey, 1972, 15) from an epistemological perspective, its status is not changing – it still is “merely” expert opinion. Hence, it is advisable that the findings of consensus methods in nursing education are further explored, discussed, deliberated and put into theoretical context before being implemented into practice. The reporting of

consensus methods should allow critical considerations with regards to how and why participants were selected, how the initial questionnaire was constructed, whether participants were provided with feedback between rounds that may have affected their views and how consensus was achieved. It is only when this information is available that it is possible to determine if particular interests of the participants led to consensus and what the political consequences could be if the consensus would be implemented. This is particularly significant with regard to the development of nursing curricula and the definition of competencies or what competencies nursing education should be based upon. In these cases, it is necessary to also examine how nursing education should be defined from a theoretical standpoint. Researchers employing consensus methods should question whether their results may lead to a narrow description of what nursing education should comprise and to then explore the consequences of implementing their results in specific political contexts. Certainly, caution should be exercised in relying solely on consensus methods as an alternative method or replacement for critically engaged dialogue and democratic debates about how to define nursing education.

5 Conclusion

Our findings are concerning if interpreted within the context of the structural critiques mentioned in the background section of our paper. These critiques question the validity of the method and our findings lend support to these critiques. If consensus methods should continue to inform best practices in nursing education, they must be rigorous in design. As such, the findings from this project have implications for all researchers using consensus methods both within the nursing profession, within healthcare and beyond. It would be interesting to analyze in more detail why consensus

methods are used in nursing education. Consensus methods may enable the realization of political agendas and their increased use could be due to their ability to serve as legitimization for political decision making. This would be particularly interesting to analyze in the context of the transformation of health care systems and academic education in many Western countries. Nevertheless, we are convinced that consensus methods have their place in nursing education, especially if it is about unravelling the “tacit” knowledge of experts. Consensus methods are certainly useful in the context of technical, biomedical dimensions of nursing education but they should not replace critically engaged dialogue and democratic debates about what nursing education should comprise beyond technical expertise.

Medline (Ovid), Embase (Ovid), AMED (Ovid),
ERIC (Ovid) and CINAHL (Ovid)
(Search terms: “Delphi”, “RAND”, “nominal group”, “consensus group methods”
AND “nursing education”). Time limit: 2004-2014

1,569 articles identified

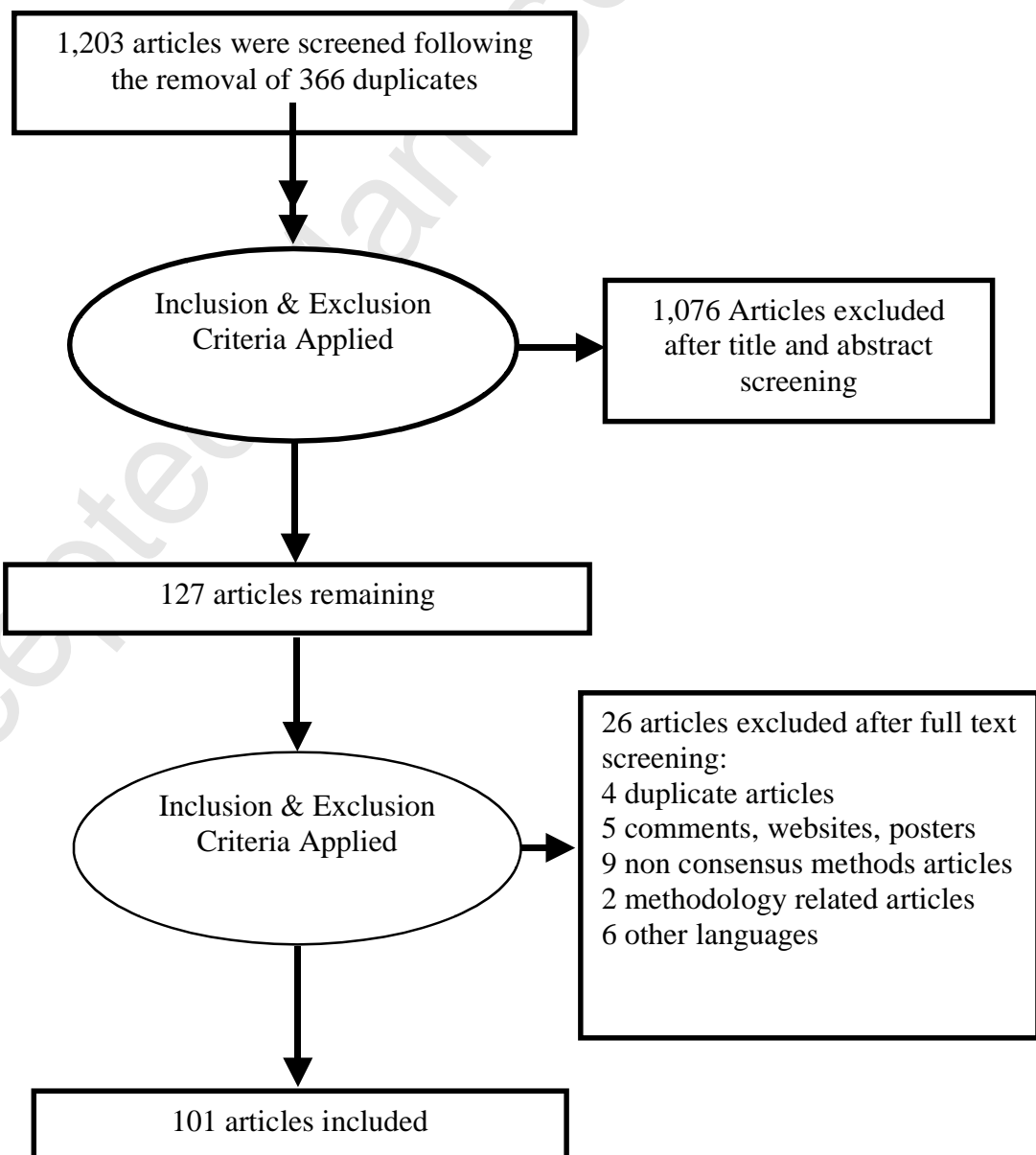


Diagram 1

Accepted Manuscript

References

- Black, J., Allen, D., Redfern, L., Muzio, L., Rushowick, B., Balaski, B., Martens, P., Crawford, M., Conlin-Saindon, K., Chapman, L., Gautreau, G., Brennan, M., Gosbee, B., Kelly, C., Round, B., 2008. Competencies in the context of entry-level registered nurse practice: a collaborative project in Canada. *International nursing review* (55), 171-178.
- Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., Albeti, C., 2011. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS ONE* 6, e20476.
- Campbell, S.M., Cantrill, J.A., 2001. Consensus methods in prescribing research. *Journal of clinical pharmacy and therapeutics* (26), 5-14.
- Carper, B.A., 1978. Fundamental patterns of knowing in nursing. *Advances in Nursing Science* 1 (1), 13-23.
- College of Nurses of O., 2014. Competencies for entry-level Registered Nurses practice. College of Nurses of Ontario, Toronto.
- Crisp, J., Pelletier, D., Duffield, C., Adams, A., Nagy, S., 1997. The Delphi Method? *Nursing research* 46, 116-118.
- Dajani, J.S., Sincoff, M.Z., Talley, W.K., 1979. Stability and agreement criteria for the termination of Delphi studies. *Technological Forecasting and Social Change* 13 (1), 83-90.
- Dalkey, N.C., 1972. The Delphi method: An experimental study of group opinion. In: Dalkey, N.C., Rourke, D. L., Lewis, R., Snyder, D. (Ed.), *Studies in the quality of life: Delphi and decision making*. Lexington Books, Lexington, MA, pp. 13-54.
- Diamond, I.R., Grant, R.C., Feldman, B.M., Pencharz, P.B., Ling, S.C., Moore, A.M., Wales, P.W., 2014. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology* 67 (4), 401-409.
- DiCenso, A., Guyatt, G., Ciliska, D., 2005. *Evidence-Based Nursing. A Guide to Clinical Practice*. Mosby, St. Louis.
- Evans, K., 1995. Competence and Citizenship: Towards a Complementary Model for Times of Critical Social Change. *British Journal of Education & Work* 8 (2), 14-27.
- Fink, A., Kosecoff, J., Chassin, M., Brook, R.H., 1984. Consensus Methods: Characteristics and Guidelines for Use. *American Journal of Public Health* 74 (9), 979-983.
- Fischer, M., 2015. Fit for the Future? A New Approach in the Debate about What Makes healthcare Systems Really sustainable. *Sustainability* (7), 312.
- Goodman, C.M., 1987. The Delphi technique: a critique. *Journal of Advanced Nursing* 12, 729-734.

- Granovskiy, B., Gold, J.M., Sumpter, D.J.T., Goldstone, R.L., 2015. Integration of Social Information by Human Groups. *Topics in Cognitive Science* 7 (3), 469-493.
- Greco, M., Zangrillo, A., Mucchetti, M., Nobile, L., Landoni, P., Bellomo, R., Landoni, G., 2015. Democracy-Based Consensus in Medicine. *Journal of Cardiothoracic and Vascular Anesthesia* 29 (2), 506-509.
- Hasson, F., Keeney, S., McKenna, H., 2000. Research guidelines for the Delphi survey technique. *Journal of advanced nursing* 32, 1008-1015.
- Haug, C., 2015. What is consensus and how is it achieved in meetings? Four types of consensus decision-making. In: Allen, J.A., Lehmann-Willenbrock, N., Rogelberg, S.G. (Eds.), *The Cambridge Handbook of Meeting Science*. Cambridge University Press, Cambridge, pp. 370-400.
- Helmer, O., Rescher, N., 1959. On the Epistemology of the Inexact Sciences. *Management Science* 6 (1), 25-52.
- Humphrey-Murto S, U.L., Consalves C, Varpio L, Foth T, Falconer M, Lynch S, Wood T, 2014. The Use of Consensus Group Methods in Medical Education Research. In: *Canadian Conference on Medical Education*. Ottawa, Ontario.
- Jairath, N., Weinstein, J., 1994. The Delphi methodology (Part one): A useful administrative approach. *Canadian journal of nursing administration* 7, 29-42.
- Jones, J., Hunter, D., 1995. Consensus methods for medical and health services research. *BMJ (Clinical research ed.)* 311, 376-380.
- Kea, B., Sun, B., 2015. Consensus development for healthcare professionals. *Internal and Emergency Medicine* 10 (3), 373-383.
- Keeney, S.H., F; McKenna, HP, 2001. A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies* 38 (2), 200.
- McKenna, H.P., 1994. The Delphi technique: a worthwhile research approach for nursing? *Journal of Advanced Nursing* 19 (6), 1221-1225.
- Murphy, M.K., Black, N.A., Lamping, D.L., McKee, C.M., Standerson, C.F.B., Askam, J., 1998. Consensus development methods, and their use in clinical guideline development. *Health Technol Assessment* 2.
- Naylor, C.D., 1995. Grey zones of clinical practice. Some limits to evidence-based medicine. *Lancet* 345, 840-842.
- Nguyen-Lu, N., Downey, K., Carvalho, J.A., 2015. Controversy between anesthesiologists and obstetricians on the labour ward: the Delphi method is used as a consensus-building technique. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie* 62 (3), 271-277.
- Padgett, K., Rhodes, C., Lumb, M., Morris, P., sherwin, S., Symons, J., Tate, J., Townend, K., 2014. What matters to users of services? An explorative study to promote shared decision making in health care. *Health Expectations* 17, 418-428.
- Rennie, D., 1981. Consensus Statements. *N Engl J Med* 304 (11), 665-666.
- Rowe, G., Wright, G., Bolger, F., 1991. Delphi: A reevaluation of research and theory. *Technological Forecasting and Social Change* 39 (3), 235-251.
- Sackett, D.L., 1997. *Evidence-based medicine: How to practice and teach EBM*. Curchill Livingstone, New York.

- Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W.M.C., Haynes, R.B., 2000. Evidence-based Medicine: How to Practice and Teach EBM. Churchill Livingstone, New York.
- Sackman, H., 1974. Delphi critique. Lexington Books, Lexington, Massachusetts.
- Scott, E.A., 1991. Appropriateness of cholecystectomy in the United Kingdom - a consensus panel approach. *Gut* 32, 1066-1070.
- Sinha, I.P., Smyth, R.L., Williamson, P.R., 2011. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. *PLoS Med* 8, e1000393.
- Stewart, J., 2001. Is the Delphi technique a qualitative method? *Medical education* 35, 922-923.
- Tammela, O., 2013. Applications of consensus methods in the improvement of care of paediatric patients: a step forward from a 'good guess'. *Acta Paediatrica* 102, 111-115.
- Trevelyan, E.G.R., Prof Nicola, 2015. Delphi methodology in health research: how to do it? *European Journal of Integrative Medicine* 7 (4), 428.
- von der Gracht, H.A., 2012. Consensus measurement in Delphi studies. *Technological Forecasting & Social Change* 79 (8), 1525-1536.
- Woudenberg, F., 1991. An evaluation of Delphi. *Technological Forecasting and Social Change* 40 (2), 131-150.